

METHODOLOGIES FOR IDENTIFYING RURAL AND NON-RURAL POPULATIONS

Two alternative methodologies for determining the rural and non-rural character of Alaska populations were developed during this project. The two alternative methodologies are called *Discriminant Analysis Assessment* and *Criterion-Referenced Assessment*. Each approach is described in this section, along with a detailed application of each methodology to a test set of 195 case populations. Based on the outcomes from the tests, both methodologies were found to produce similar classifications of case populations. Because of its comparative simplicity, *Criterion-Referenced Assessment* might be considered a preferred method for identifying rural and non-rural populations for federal subsistence management. The *Discriminant Analysis Assessment* also might be considered a useful approach because it provides a quantitative measure of the closeness of each community to a rural or non-rural class. The validity of either approach receives support by the similar outcomes in the test analyses.

Methodology 1. Discriminant Analysis Assessment

Discriminant analysis (also called *discriminant function analysis*) is a statistical method designed to distinguish between one or more groups. It is a type of multivariate modeling that relies on correlation and multiple regression of variables measured with interval data (or near interval data). For this assessment, we have used the discriminant analysis program in the *Statistical Package for the Social Sciences (SPSS)*, a commonly-used social science statistical software.

Using discriminant analysis, variables are identified that distinguish between groups of cases (the groups are called the *dependents*). The variables found to statistically distinguish between groups are called *discriminating variables* (also called *predictors* or *independents*). The discriminating variables are statistically combined into numeric equations called *discriminant functions* (also called *canonical roots*). A discriminant function can be treated as a *criterion* for distinguishing among groups. Depending upon the data and cases analyzed, one criterion or several criteria may emerge through discriminant analysis for distinguishing between groups.

The discriminant functions can be used to classify uncertain cases into the groups of cases. For case populations, discriminating variables are measured and the values entered into the numeric equations (the discriminant functions). The case population's score indicates its group. The case is classified with the group to which its numeric value is closest to the numeric *centroid* of each group. The nearness of the case population's score to the group's centroid indicates the clarity or ambiguity of the classification.

Discriminant analysis was conducted with a selection of Alaska populations comprising the set of cases, while the dependents (groups) were "rural" or "non-rural." In this test of

the methodology, case populations were included in the analysis if they had greater than 49 people, country food harvests of less than 1,000 lbs per capita, and information on discriminating variables. There were 195 populations meeting the selection criteria in the data set. The selection was done to reduce the number of potential case outliers (cases with very small or large values), which can confound statistical correlations. Very small populations (<50 people) are likely to be classed as “rural” under most standards and may display unusual traits linked to their size. Populations with exceptionally large annual productions of country foods (such as those with harvests greater than 1,000 lbs per year) also are likely to be classed as “rural,” regardless of any other characteristic they may display.

Two *discriminating variables* were identified in analysis: (1) the annual per capita harvest of country foods by the population – log transformed (LGPCAP3); and (2) the density of population to a standard area – log transformed (LGDEN30). These variables are measures of primary rural concepts, as described above. Before analysis, log transformations of per capita harvests and density were made because the frequency distributions of their values were asymmetric (skewed) and the two variables appeared to have a curvilinear relationship. Discriminant analysis works better with normally distributed values and linear relationships. The log transformations produce values with greater symmetry and linearity.

Fourteen separate discriminant function analyses were conducted with the two discriminating variables to test the methodology. Each analysis was conducted with slightly different starting conditions, variable measurements, or case population definitions. A summary of the fourteen separate analyses is presented as Appendix B. A discussion of the fourteen analyses also is presented in the appendix. In this section, the “best analysis” from these fourteen runs is presented in detail (shown as *Run A* in Appendix B Table). This best analysis discriminated the greatest amount of the variability among case values, as measured by the *canonical correlation*. The model discriminated 82.3% of the variability in the two discriminating variables among the 195 case populations, which is a very high value (a canonical correlation of .907). It was the highest among the fourteen analyses. The analysis provided excellent discrimination between groups, as described below.

In discriminant analysis, initial groupings of cases are advanced to focus the analysis on appropriate discriminating variables and separation points. In the “best analysis,” cases were assigned to initial groups based on classifications of the Federal Subsistence Board and the State Joint Board of Fisheries and Game. Case populations classed as “rural” by both boards were initially labeled “rural” (132 populations). Case populations labeled as “non-rural” by both boards were initially labeled “non-rural” (54 populations). Case populations for which the federal and state board classifications differed were left unclassified (9 populations).

Initial groupings based on *population thresholds* (“rural” <2,500 people, “unclassified” 2,500-7,000 people, and “non-rural” >7,000 people) also were examined in seven runs (Runs B and 7-12 in Appendix B). Runs with initial groupings based on population

thresholds produced discriminant functions with reduced discriminating capabilities (shown by canonical correlations ranging from .863 to .888). As these runs performed less well in separating rural and non-rural groups, this approach was not chosen as a “best analysis.”

In the “best analysis,” density was measured using the weighted populations within a *30-mile standard area*. Other runs using 10-mile and 20-mile standard areas for measuring density produced only slightly different outcomes, as discussed in Appendix B. Country food harvests for case populations with dual harvest estimates (certain Kenai Peninsula Borough populations) were estimated with the average of two per capita values (Country Food Harvest = Community Profile Database estimate + Harvest Ticket/Permit Record estimate / 2). Outcomes of runs separately using CPDB harvest estimates or using the Harvest Ticket/Permit Record estimates were also run, presented in Appendix B. The following discussion describes the “best analysis” in detail.

In discriminant analysis, the mean values of discriminating variables are calculated. In the “best analysis,” the mean values of discriminating variables differed substantially between the initial groupings of cases, as shown in the following table entitled *Group Statistics* (1.00 = non-rural group; 2.00 = rural group; Total = pooled cases).

Group Statistics

TESTRUR		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
1	LGDEN30	2.7938	.59461	54	54.000
	LGPCAP3	1.3436	.25789	54	54.000
2	LGDEN30	1.0549	.49540	132	132.000
	LGPCAP3	2.5142	.27518	132	132.000
Total	LGDEN30	1.5597	.94945	186	186.000
	LGPCAP3	2.1744	.59711	186	186.000

The group of 54 case populations initially labeled “non-rural” has a mean country food production (LGPCAP3) of 1.3436 (equivalent to 22.1 lbs per capita), while the group of 132 case populations initially labeled “rural” has a mean of 2.5142 (equivalent to 326.7 lbs per capita). The non-rural group has a mean density (LGDEN30) of 2.7938 (equivalent to 6.2 weighted persons per sq mi), while the rural group has a mean of 1.0549 (equivalent to 0.1 weighted persons per sq mi). The differences between rural and non-rural groups are statistically significant (sig. < .000). The nine unclassified cases do not figure into the discriminant analysis at this stage.

The *covariances* of discriminating variables are calculated for the pool of 195 case populations, and for the rural and non-rural groups separately (1.00 = non-rural group; 2.00 = rural group; Total = pooled cases), as shown in the following two matrices. Covariance is the sum of squared distances of each case population from the group mean (shown in the Group Statistics table, above). For example, the covariance of LGPCAP3

with LGDEN30 is -0.0507 in the pooled group. The *correlations* of variables with one another are also shown in the pooled matrix.

Pooled Within-Groups Matrices^a

		LGDEN30	LGPCAP3
Covariance	LGDEN30	.277	-5.07E-02
	LGPCAP3	-5.07E-02	7.307E-02
Correlation	LGDEN30	1.000	-.357
	LGPCAP3	-.357	1.000

a. The covariance matrix has 184 degrees of freedom.

Covariance Matrices^a

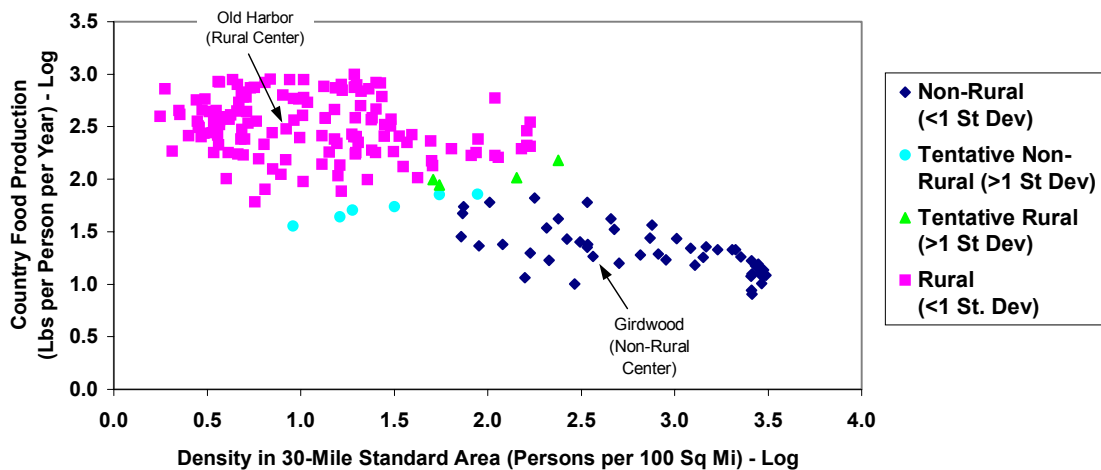
TESTRUR		LGDEN30	LGPCAP3
1	LGDEN30	.354	-.107
	LGPCAP3	-.107	6.651E-02
2	LGDEN30	.245	-2.78E-02
	LGPCAP3	-2.78E-02	7.573E-02
Total	LGDEN30	.901	-.472
	LGPCAP3	-.472	.357

a. The total covariance matrix has 185 degrees of freedom.

In this analysis, the two *discriminating variables* were inserted into the equation together (an alternative for analysis with more than two variables is to use a step-wise insertion method, where the best-distinguishing variables are inserted in order). The analysis assesses the extent to which the two variables (country food production and density), considered jointly, separate cases into two distinct groups (rural and non-rural).

One way to visually represent the relationship between country food production (LGPCAP3) and density (LGDEN30) for the 195 case populations is shown in Fig. 13. Each symbol is a case population. There is a clear linear relationship between the two variables (the correlation coefficient is -0.357, as shown in the *Pooled Within Group Matrices* table above). Also, the scatter plot illustrates a cluster of case populations toward the upper left quadrant (rural) and a cluster toward the lower right quadrant (non-rural) with a noticeable separation. The two variables appear to separate case populations into fairly distinct clusters in the upper left and lower right quadrants. The lower left and upper right quadrants are empty of cases. Although cases with those values can be imagined (that is, low density-low production populations and high density-high production populations), they apparently are not common in Alaska (at least as shown by this test sample of 195 Alaska populations).

Fig. 13. Alaska Populations (N = 195) Categorized into Rural or Non-Rural Groups by two Primary Factors (Density and Country Food Production), With Standard Deviations from Group Centers



The relationships displayed in Fig. 13 are expected, based on the theoretical propositions in *Rural Measures*. That is, it is expected that relatively lower levels of country food production occur in higher density populations, while relatively higher levels of country food production occur in lower density populations.

The next step of the discriminant analysis is to calculate the *discriminant function*, which is a latent variable (statistically-constructed variable) created as a linear combination of discriminating variables, taking the form of an equation $L = b_1x_1 + b_2x_2 + \dots b_nx_n + c$. In this equation, the b 's are *discriminant coefficients*, the x 's represent the values of discriminating variables, and c is some constant number. The discriminant function equation is analogous to a multiple regression equation, except that the b 's are discriminant coefficients. The discriminant function is estimated using ordinary least-squares.

The discriminant function can be described through several summary statistics. The *canonical correlation* indicates the percent of variation in the *dependent variable* (rural or non-rural) discriminated by the *independent variables* (country food production, population density) in discriminant analysis. In this case, a square of the canonical correlation (0.907), shown in the *Eigenvalues* table below, indicates that 82.3% of the variation is discriminated (a high value).

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	4.646 ^a	100.0	100.0	.907

a. First 1 canonical discriminant functions were used in the analysis.

The *standardized coefficients* indicate the relative importance of each discriminating variable in predicting the dependent. In this example, country food production (LGPCAP3) at 0.764 is a somewhat more important discriminating variable than population density (LGDEN30) at -0.427, as shown in the Standardized Canonical Discriminant Function Coefficients table. But both are important contributors to the discriminating function.

Standardized Canonical Discriminant Function Coefficients

	Function
	1
LGDEN30	-.427
LGPCAP3	.764

The *canonical discriminant function coefficients* are the values inserted for b 's in the discriminant function equation. In this example, the discriminant function equation is $L = 2.828(LGPCAP3) - .812(LGDEN30) - 4.882$, as indicated in the *Canonical Discriminant Function Coefficients* table.

Canonical Discriminant Function Coefficients

	Function
	1
LGDEN30	-.812
LGPCAP3	2.828
(Constant)	-4.882

Unstandardized coefficients

By entering the values for each case population into this equation, one calculates a case population's score on the discriminant function. A listing of each case population's score is shown in the following table, *Outcome Scores and Classifications from Discriminant Analysis Assessment*.

Outcome Scores and Classifications from Discriminant Analysis Assessment

	Population	Density (Log)	Country Food Production (Log)	Initial Classification	Discriminant Analysis Outcome Classification	Score	Distance From Non- Rural Center	Distance From Rural Center
1	Girdwood	2.56	1.26	Non-Rural	Non-Rural	-3.39	0.04	4.76
2	Nikiski	2.33	1.23	Non-Rural	Non-Rural	-3.31	0.04	4.68
3	Eagle River	3.01	1.44	Non-Rural	Non-Rural	-3.26	0.09	4.64
4	Fort Wainwright	2.82	1.28	Non-Rural	Non-Rural	-3.55	0.20	4.92
5	North Pole Area	2.87	1.44	Non-Rural	Non-Rural	-3.14	0.21	4.51
6	Upper OMalley	3.09	1.34	Non-Rural	Non-Rural	-3.59	0.24	4.96
7	Eielson AFB	2.53	1.35	Non-Rural	Non-Rural	-3.11	0.24	4.48
8	Southwest Fairbanks	2.91	1.29	Non-Rural	Non-Rural	-3.61	0.26	4.98
9	Rabbit Creek	3.17	1.35	Non-Rural	Non-Rural	-3.62	0.27	4.99
10	Houston	2.20	1.06	Non-Rural	Non-Rural	-3.66	0.31	5.03
11	Wasilla (group)	2.54	1.38	Non-Rural	Non-Rural	-3.03	0.32	4.40
12	Northwest Fairbanks	2.70	1.20	Non-Rural	Non-Rural	-3.68	0.33	5.05
13	Big Lake	2.23	1.30	Non-Rural	Non-Rural	-3.02	0.33	4.39
14	Coastal Refuge	3.23	1.33	Non-Rural	Non-Rural	-3.75	0.39	5.12
15	Juneau City and Borough	2.49	1.40	Non-Rural	Non-Rural	-2.94	0.41	4.31
16	Central Fairbanks	2.95	1.23	Non-Rural	Non-Rural	-3.80	0.44	5.17
17	Lower OMalley-Cambell	3.31	1.33	Non-Rural	Non-Rural	-3.81	0.46	5.18
18	OMalley	3.33	1.33	Non-Rural	Non-Rural	-3.82	0.47	5.20
19	Elmendorf	3.15	1.26	Non-Rural	Non-Rural	-3.89	0.54	5.26
20	Palmer (group)	2.42	1.43	Non-Rural	Non-Rural	-2.81	0.55	4.18
21	Chugiak	2.88	1.56	Non-Rural	Non-Rural	-2.80	0.56	4.17
22	Northeast Fairbanks	2.68	1.52	Non-Rural	Non-Rural	-2.75	0.60	4.13
23	Airport	3.35	1.26	Non-Rural	Non-Rural	-4.04	0.69	5.41
24	Sutton-Alpine	2.08	1.38	Non-Rural	Non-Rural	-2.67	0.69	4.04
25	North Fairbanks	2.47	1.00	Non-Rural	Non-Rural	-4.05	0.70	5.42
26	Fort Richardson	3.11	1.18	Non-Rural	Non-Rural	-4.07	0.72	5.44
27	Willow (group)	1.95	1.37	Non-Rural	Non-Rural	-2.60	0.75	3.98
28	Muldoon	3.41	1.22	Non-Rural	Non-Rural	-4.20	0.85	5.57
29	Eklutna	2.66	1.62	Non-Rural	Non-Rural	-2.45	0.90	3.82
30	Ketchikan	2.32	1.54	Non-Rural	Non-Rural	-2.42	0.93	3.79
31	Campbell Creek	3.45	1.19	Non-Rural	Non-Rural	-4.32	0.97	5.69
32	Little Campbell Creek	3.43	1.18	Non-Rural	Non-Rural	-4.34	0.99	5.71
33	Seward (group)	1.86	1.45	Non-Rural	Non-Rural	-2.28	1.07	3.65
34	Northfork	3.46	1.14	Non-Rural	Non-Rural	-4.48	1.13	5.85
35	Soldotna (group)	2.38	1.62	Non-Rural	Non-Rural	-2.22	1.13	3.59
36	Midtown	3.47	1.13	Non-Rural	Non-Rural	-4.50	1.15	5.87
37	Delaney Lake	3.46	1.11	Non-Rural	Non-Rural	-4.54	1.19	5.91
38	Spennard	3.41	1.10	Non-Rural	Non-Rural	-4.54	1.19	5.91
39	Ship Creek	3.41	1.08	Non-Rural	Non-Rural	-4.60	1.25	5.97
40	University	3.47	1.09	Non-Rural	Non-Rural	-4.62	1.27	6.00
41	MidFork-RusJack	3.46	1.08	Non-Rural	Non-Rural	-4.63	1.28	6.00
42	Russian Jack	3.48	1.08	Non-Rural	Non-Rural	-4.64	1.29	6.01
43	Lake Otis	3.49	1.09	Non-Rural	Non-Rural	-4.64	1.29	6.02
44	Kenai	2.53	1.78	Non-Rural	Non-Rural	-1.91	1.44	3.28
45	Merrill Field	3.46	1.01	Non-Rural	Non-Rural	-4.85	1.50	6.22
46	Avenue Fifteen	3.41	.94	Non-Rural	Non-Rural	-4.98	1.63	6.36
47	Salcha-Harding	1.86	1.68	Non-Rural	Non-Rural	-1.66	1.69	3.03
48	Downtown	3.41	.91	Non-Rural	Non-Rural	-5.09	1.74	6.47
49	Homer	2.25	1.82	Non-Rural	Non-Rural	-1.56	1.79	2.93
50	Kasilof (group)	2.01	1.78	Non-Rural	Non-Rural	-1.48	1.87	2.85
51	Anchor Point (group)	1.87	1.74	Non-Rural	Non-Rural	-1.48	1.87	2.85
52	Glacier View CDP	.96	1.55	Rural	Tentative Non-Rural	-1.27	2.09	2.64
53	Moose Pass (group)	1.21	1.64	Non-Rural	Tentative Non-Rural	-1.22	2.14	2.59
54	Fritz Creek CDP	1.94	1.86	Uncertain	Tentative Non-Rural	-1.21	2.14	2.58
55	Talkeetna	1.50	1.74	Uncertain	Tentative Non-Rural	-1.18	2.17	2.55
56	Trapper Creek	1.28	1.71	Uncertain	Tentative Non-Rural	-1.10	2.26	2.47
57	North Fork Road	1.74	1.85	Uncertain	Tentative Non-Rural	-1.06	2.29	2.43
58	Old Harbor	.92	2.48	Rural	Rural	1.37	0.00	4.73
59	Manokotak	1.29	2.58	Rural	Rural	1.38	0.01	4.73
60	Coffman Cove	.85	2.44	Rural	Rural	1.33	0.04	4.68
61	Yakutat	1.38	2.59	Rural	Rural	1.31	0.06	4.66
62	Naukat Bay	.68	2.38	Rural	Rural	1.31	0.07	4.66
63	Kotzebue	2.04	2.77	Rural	Rural	1.30	0.07	4.66
64	McKinley Park Village	.70	2.38	Rural	Rural	1.29	0.08	4.64
65	Whale Pass	.31	2.27	Rural	Rural	1.28	0.10	4.63

Outcome Scores and Classifications from Discriminant Analysis Assessment (p. 2)

	Population	Density (Log)	Country Food Production (Log)	Initial Classification	Discriminant Analysis Outcome Classification	Score	Distance From Non- Rural Center	Distance From Rural Center
66	Galena	1.38	2.57	Rural	Rural	1.25	0.12	4.61
67	Aleknagik	1.13	2.58	Rural	Rural	1.49	0.12	4.84
68	Chistochina	.56	2.42	Rural	Rural	1.50	0.13	4.86
69	Chignik Lagoon	.56	2.33	Rural	Rural	1.24	0.13	4.59
70	Hydaburg	1.13	2.58	Rural	Rural	1.51	0.14	4.86
71	Akutan	1.40	2.67	Rural	Rural	1.53	0.15	4.88
72	Nelson Lagoon	.47	2.40	Rural	Rural	1.54	0.17	4.89
73	South Naknek	.69	2.47	Rural	Rural	1.55	0.18	4.91
74	Hoonah	1.48	2.57	Rural	Rural	1.18	0.19	4.53
75	Clark's Point	.96	2.56	Rural	Rural	1.57	0.20	4.93
76	Northway	.53	2.44	Rural	Rural	1.60	0.23	4.95
77	Bettles-Evansville	.40	2.42	Rural	Rural	1.62	0.25	4.97
78	Cheneg Bay	.48	2.44	Rural	Rural	1.62	0.25	4.98
79	Tanacross	.99	2.40	Rural	Rural	1.09	0.28	4.44
80	Tatitlek	1.01	2.61	Rural	Rural	1.67	0.30	5.03
81	Port Lions	1.45	2.52	Rural	Rural	1.07	0.30	4.42
82	Kotlik	1.32	2.70	Rural	Rural	1.68	0.31	5.04
83	Lake Louise	.53	2.25	Rural	Rural	1.06	0.31	4.41
84	Tetlin	.80	2.33	Rural	Rural	1.05	0.32	4.41
85	Noatak	1.18	2.66	Rural	Rural	1.69	0.32	5.04
86	Tyonek	1.11	2.41	Rural	Rural	1.04	0.33	4.39
87	Chitina	.72	2.53	Rural	Rural	1.70	0.33	5.05
88	Pelican	.76	2.55	Rural	Rural	1.71	0.34	5.06
89	Klawock	1.48	2.51	Rural	Rural	1.00	0.37	4.35
90	Whitestone Logging Can	.61	2.25	Rural	Rural	0.99	0.38	4.34
91	Tenakee Springs	.57	2.52	Rural	Rural	1.78	0.41	5.13
92	Saint Paul	1.27	2.43	Rural	Rural	0.95	0.42	4.30
93	Port Alexander	.46	2.49	Rural	Rural	1.80	0.43	5.15
94	Slana	.66	2.24	Rural	Rural	0.91	0.46	4.26
95	Emmonak	1.43	2.79	Rural	Rural	1.83	0.46	5.19
96	Akhiok	.45	2.51	Rural	Rural	1.84	0.47	5.19
97	Gustavus	1.18	2.38	Rural	Rural	0.89	0.48	4.25
98	Hyder	.54	2.54	Rural	Rural	1.86	0.49	5.21
99	Nanwalek	1.28	2.40	Rural	Rural	0.88	0.50	4.23
100	Port Graham	1.30	2.40	Rural	Rural	0.86	0.51	4.21
101	Larsen Bay	.62	2.57	Rural	Rural	1.88	0.51	5.23
102	Hollis	.69	2.23	Rural	Rural	0.86	0.51	4.21
103	Port Alsworth	.57	2.56	Rural	Rural	1.89	0.52	5.24
104	Egegik	.61	2.58	Rural	Rural	1.93	0.56	5.28
105	King Salmon	1.19	2.34	Rural	Rural	0.77	0.60	4.13
106	Chignik Bay	.45	2.55	Rural	Rural	1.98	0.60	5.33
107	Pilot Point	.55	2.58	Rural	Rural	1.98	0.61	5.33
108	Perryville	.58	2.60	Rural	Rural	1.99	0.62	5.34
109	King Cove	1.45	2.41	Rural	Rural	0.75	0.62	4.10
110	Port Heiden	.62	2.61	Rural	Rural	1.99	0.62	5.34
111	Tanana	1.04	2.73	Rural	Rural	2.00	0.63	5.35
112	Chignik Lake	.71	2.65	Rural	Rural	2.02	0.65	5.38
113	Angeon	1.31	2.35	Rural	Rural	0.71	0.67	4.06
114	Tonsina	.78	2.19	Rural	Rural	0.69	0.68	4.04
115	Sand Point	1.53	2.41	Rural	Rural	0.69	0.69	4.04
116	Fort Yukon	1.32	2.84	Rural	Rural	2.06	0.69	5.41
117	Ouzinkie	1.59	2.42	Rural	Rural	0.67	0.70	4.02
118	Shageluk	.66	2.65	Rural	Rural	2.07	0.70	5.42
119	Alakanuk	1.36	2.86	Rural	Rural	2.10	0.73	5.45
120	Brevig Mission	.99	2.76	Rural	Rural	2.13	0.76	5.48
121	Huslia	1.02	2.78	Rural	Rural	2.15	0.77	5.50
122	Minto	.96	2.77	Rural	Rural	2.16	0.79	5.52
123	McGrath	1.15	2.26	Rural	Rural	0.57	0.80	3.92
124	New Stuyahok	1.22	2.85	Rural	Rural	2.17	0.80	5.52
125	Atka	.51	2.64	Rural	Rural	2.18	0.80	5.53
126	Nikolai	.55	2.65	Rural	Rural	2.18	0.80	5.53
127	Gulkana	.92	2.18	Rural	Rural	0.55	0.83	3.90
128	Mountain Village	1.43	2.91	Rural	Rural	2.20	0.83	5.55
129	Wainwright	1.29	2.88	Rural	Rural	2.20	0.83	5.56
130	False Pass	.36	2.62	Rural	Rural	2.23	0.86	5.58

Outcome Scores and Classifications from Discriminant Analysis Assessment (p. 3)

		Density	Country Food	Initial	Discriminant Analysis		Distance	Distance
	Population	(Log)	Production	Classification	Outcome Classification	Score	From Non- Rural Center	From Rural Center
131	Quinhagak	1.29	2.89	Rural	Rural	2.23	0.86	5.58
132	Sitka Tribe	2.23	2.54	Rural	Rural	0.50	0.87	3.86
133	Kwethluk	1.40	2.92	Rural	Rural	2.24	0.87	5.59
134	Chickaloon	1.57	2.35	Rural	Rural	0.49	0.88	3.84
135	Beaver	.47	2.66	Rural	Rural	2.26	0.88	5.61
136	Shishmaref	1.30	2.90	Rural	Rural	2.26	0.89	5.61
137	Pedro Bay	.25	2.60	Rural	Rural	2.27	0.89	5.62
138	Nuiqsut	1.19	2.87	Rural	Rural	2.27	0.90	5.62
139	Allakaket/Alatna	.67	2.73	Rural	Rural	2.30	0.93	5.65
140	Thorne Bay	1.29	2.25	Rural	Rural	0.44	0.93	3.79
141	Holy Cross	.90	2.80	Rural	Rural	2.31	0.94	5.66
142	Craig	1.69	2.37	Rural	Rural	0.43	0.94	3.78
143	Naknek	1.38	2.27	Rural	Rural	0.43	0.94	3.78
144	Copper Center	1.29	2.24	Rural	Rural	0.41	0.96	3.76
145	Port Protection	.35	2.65	Rural	Rural	2.34	0.97	5.69
146	Nunapitchuk	1.22	2.90	Rural	Rural	2.34	0.97	5.69
147	Kivalina	1.13	2.88	Rural	Rural	2.35	0.98	5.70
148	Mentasta Lake	.85	2.10	Rural	Rural	0.36	1.01	3.71
149	Kake	1.40	2.25	Rural	Rural	0.35	1.02	3.70
150	Golovin	.71	2.78	Rural	Rural	2.41	1.04	5.76
151	Klukwan	.69	2.78	Rural	Rural	2.43	1.06	5.78
152	Seldovia	1.50	2.26	Rural	Rural	0.30	1.07	3.66
153	Skwentna (group)	.60	2.00	Rural	Rural	0.30	1.08	3.65
154	Barrow	2.21	2.46	Rural	Rural	0.28	1.09	3.63
155	Dillingham	1.95	2.38	Rural	Rural	0.28	1.09	3.63
156	Anderson	1.11	2.14	Rural	Rural	0.28	1.10	3.63
157	Stevens Village	.49	2.76	Rural	Rural	2.53	1.16	5.88
158	Hughes	.44	2.75	Rural	Rural	2.55	1.18	5.90
159	Stebbins	1.29	3.00	Rural	Rural	2.55	1.18	5.90
160	Deering	.68	2.83	Rural	Rural	2.56	1.19	5.91
161	Cantwell	.90	2.05	Rural	Rural	0.18	1.19	3.53
162	Kenny Lake	1.21	2.13	Rural	Rural	0.17	1.20	3.52
163	Haines	1.81	2.29	Rural	Rural	0.13	1.24	3.48
164	Kaktovik	1.02	2.95	Rural	Rural	2.63	1.26	5.98
165	Newhalen	.75	2.87	Rural	Rural	2.63	1.26	5.98
166	Wales	.73	2.87	Rural	Rural	2.65	1.27	6.00
167	Point Lay	.94	2.95	Rural	Rural	2.69	1.32	6.05
168	Koliganek	.81	2.92	Rural	Rural	2.72	1.35	6.07
169	Grayling	.84	2.95	Rural	Rural	2.78	1.41	6.14
170	Ekwok	.66	2.90	Rural	Rural	2.78	1.41	6.14
171	Cordova	1.94	2.25	Rural	Rural	-0.09	1.46	3.27
172	Saxman	2.21	2.32	Rural	Rural	-0.11	1.48	3.25
173	Gakona	1.01	1.98	Rural	Rural	-0.11	1.48	3.24
174	Tazlina	1.20	2.03	Rural	Rural	-0.11	1.48	3.24
175	Tok	1.70	2.17	Rural	Rural	-0.12	1.49	3.24
176	Healy	1.55	2.12	Rural	Rural	-0.14	1.51	3.21
177	Wrangell	1.91	2.22	Rural	Rural	-0.15	1.52	3.21
178	Sitka	2.23	2.31	Rural	Rural	-0.15	1.53	3.20
179	Whittier	.81	1.90	Rural	Rural	-0.16	1.53	3.19
180	Unalaska	2.18	2.29	Rural	Rural	-0.18	1.55	3.17
181	Anvik	.57	2.93	Rural	Rural	2.93	1.56	6.28
182	Levelock	.64	2.95	Rural	Rural	2.93	1.56	6.29
183	Iliamna	.56	2.93	Rural	Rural	2.95	1.57	6.30
184	Igiugig	.27	2.86	Rural	Rural	2.98	1.61	6.34
185	Kodiak Road	2.04	2.23	Rural	Rural	-0.24	1.62	3.11
186	Ninilchik	1.71	2.13	Uncertain	Rural	-0.25	1.62	3.10
187	Petersburg	2.06	2.21	Rural	Rural	-0.31	1.68	3.04
188	Glennallen	1.36	2.00	Rural	Rural	-0.33	1.71	3.02
189	Hope	.75	1.79	Uncertain	Rural	-0.45	1.82	2.91
190	Voznesenka	1.63	2.01	Uncertain	Rural	-0.51	1.88	2.84
191	Cooper Landing	1.22	1.89	Uncertain	Rural	-0.53	1.90	2.82
192	Clam Gulch	1.71	2.00	Non-Rural	Tentative Rural	-0.62	2.00	2.73
193	Kodiak City	2.38	2.18	Rural	Tentative Rural	-0.65	2.02	2.70
194	Nikolaevsk	1.74	1.95	Uncertain	Tentative Rural	-0.79	2.16	2.56
195	Valdez	2.16	2.01	Non-Rural	Tentative Rural	-0.94	2.31	2.42

The mean value for the discriminant scores for each group (rural or non-rural) is called a *group centroid*. In this example, the group centroid is -3.352 for the non-rural group and 1.371 for the rural group, as shown in the *Functions at Group Centroids* table.

Functions at Group Centroids

	Function
TESTRUR	1
1	-3.352
2	1.371

Unstandardized canonical discriminant functions evaluated at group means

The group centroids can be used to calculate a threshold between rural and non-rural groups. The mean of the group centroids is used to define the threshold (*cutoff*) between groups. In this example, the cutoff is -0.9905 , mid-way between centroids. If a discriminant function score is above the cutoff, the case population is nearer the rural group. If a score is below the cutoff, the population is nearer in the non-rural group.

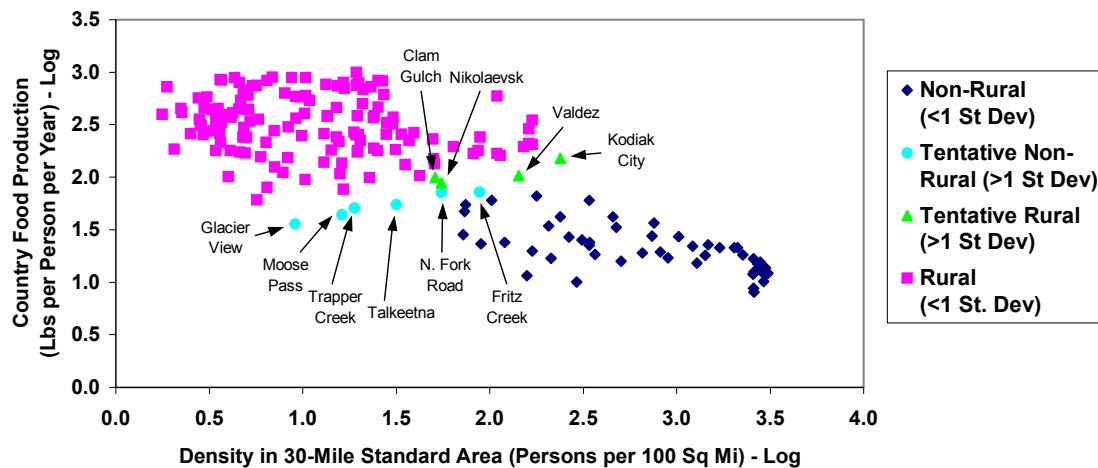
When used to classify cases, the discriminant function serves as a *criterion* for distinguishing rural and non-rural populations. In this example, the criterion is composed of a combination of two variables, country food production and density within a 30-mile standard area. As stated above, to classify case populations, the values of each case are inserted into the discriminant function equation to calculate the case population's score. The score can be compared with the threshold value separating groups, and also with the centroid value of each group. The distance between the score and the centroid is represented by a measure called the Mahalanobis distance, which is equivalent to a z-score for a normal population distribution. A Mahalanobis distance of 1.96 represents one standard deviation away from the group's center, within which 95% of cases fall under a normal curve.

In the *Outcome Scores and Classifications from Discriminant Analysis Assessment* table above, outcome scores and classifications are made for each case population. The classifications are made with the following rules. Cases with scores closer than one standard deviation from a group's centroid are classified into that group. Cases with scores greater than one standard deviation from a group centroid are *tentatively classified* in the group with the closest centroid.

Using these rules to assess the 195 case populations, 185 cases (95%) were categorized with a fair degree of certainty, defined as case scores closer than one standard deviation from the center of each group. Of these, 134 cases were categorized as "rural" and 51 cases were categorized as "non-rural." Ten case populations were given tentative classifications, defined as having scores greater than one standard deviation from the center of each group. Of the tentative cases, four were tentatively classified as "rural" and six were tentatively classified as "non-rural," representing the group to which its score was closest. A graphic depiction of these classifications is shown in Fig. 14, where cases are color-coded by the degree of certainty of their classification. As shown in Fig.

14, the tentative cases included eight populations along roads in a rural/non-rural fringe (Clam Gulch, Fritz Creek, Glacier View, Nikolaevsk CDP, North Fork Road, Talkeetna, and Trapper Creek) and two mid-sized, geographically separate populations (Kodiak City and Valdez).

Fig. 14. Alaska Populations (N = 195) Categorized into Rural or Non-Rural Groups by two Primary Factors (Density and Country Food Production), With Tentative Case Classifications Identified



A second assessment step could be taken for reviewing the classification of tentative cases, using a set of *ancillary variables* (such as other extensive land uses, specialized production, and other rural factors), drawn from existing data sets, case method materials, public comment, and additional data collection. Examples of ancillary variables are provided in the next section on *Criterion-Referenced Assessment*.

In a two-variable model, the threshold separating cases can be represented as a line, illustrated in Fig. 15. The line is defined by combinations of values of the two variables corresponding to the midpoint between group centroids. Cases to the left side of the line are closer to the “rural” group, while cases to the right side are closer to the “non-rural” group. In Fig. 15, the distances representing one standard deviation from each group’s center are also depicted as lines. Cases with scores falling within this area of the graph (greater than one standard deviation) received tentative rural or non-rural classifications, following the above classification rules.

The classification of cases uses the log transformed values of the primary discriminants – *country food production* and *density*. Transforming the discriminant functions to non-log values is a relatively simple mathematical step, allowing for another interpretation of the model. Fig. 16 provides the discriminant analysis outcome with non-log values. It shows how a case population would be classified given any combination of country food production or density (weighted population within a 30-mile distance). Cases falling above the line $y = 23.774x^{0.2874}$ are “rural,” while cases falling below are “non-rural.”

The lines $y = 17.147x^{0.2874}$ and $y = 32.953x^{0.2874}$ identify the certainty of the classifications, with cases falling between them classified as “tentative.”

Fig. 15. Alaska Populations (N = 195) Categorized into Rural or Non-Rural Groups by two Primary Factors (Density and Country Food Production), With Threshold Lines at One Standard Deviation from Each Group Center

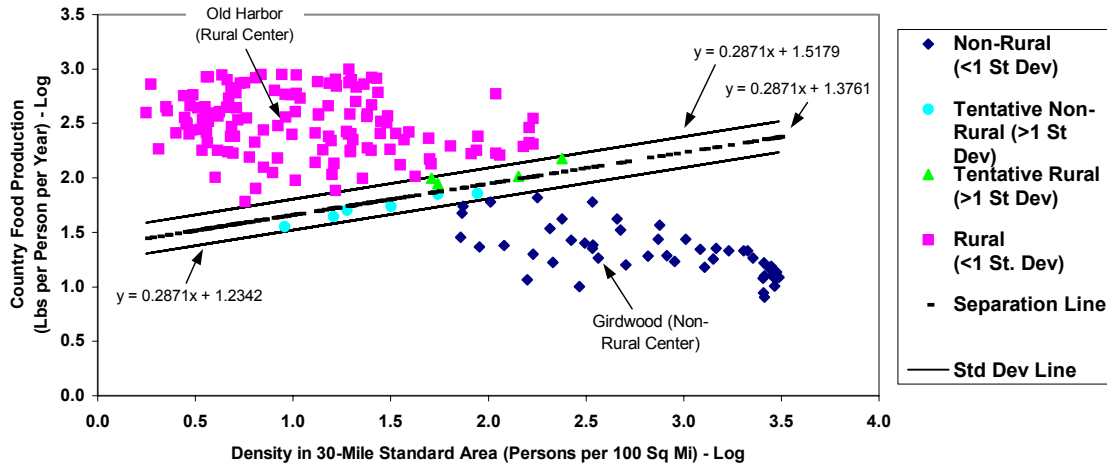
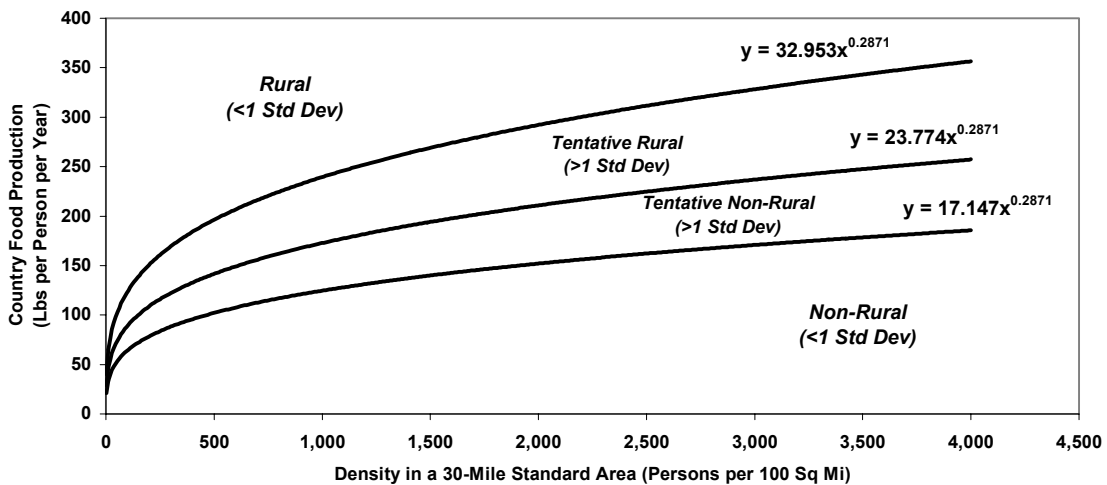
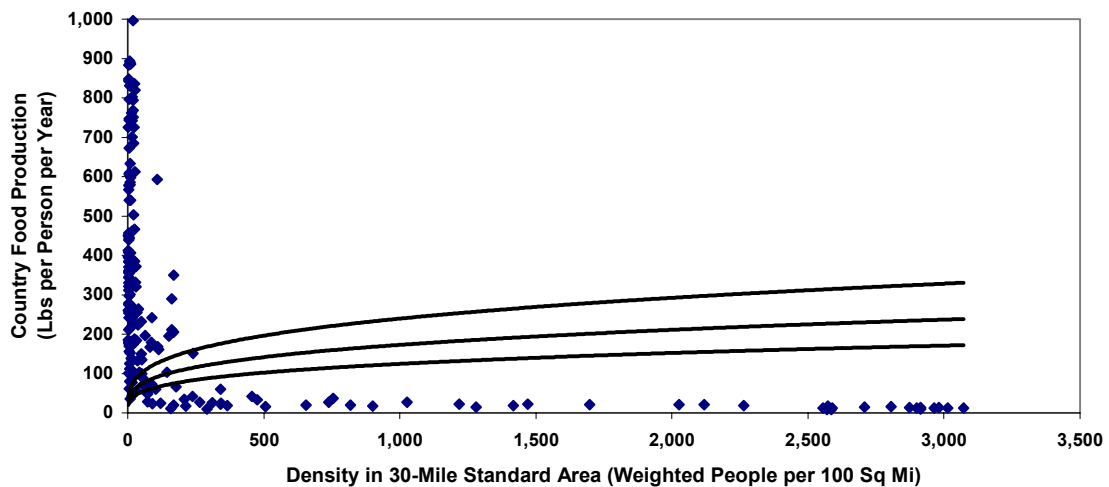


Fig. 16. General Model for Categorizing Alaska Populations into Rural or Non-Rural Groups by Two Primary Factors (Density and Country Food Production)



The actual dispersion of the 195 case populations in relation to the general classification model is shown as Fig. 17. The scatter of cases clearly reflects the bifurcated character of Alaska populations. Most of the chart (where cases might appear) is empty. Case populations appear either scattered along the y-axis (lower densities and substantial country food harvests) or along the x-axis (higher densities and insubstantial country food harvests). Ambiguous cases are relatively few, located in the interstices of the threshold lines. The ability of the discriminant function model to successfully separate cases into rural and non-rural groups results from the basic bifurcation of Alaska populations along the two measures.

Fig. 17. Relationship of Density and Country Food Production in 195 Alaska Populations, and General Classification Model



Methodology 2. Criterion-Referenced Assessment

Criterion-Referenced Assessment is a general methodology for classifying cases into categories. Criterion-referenced assessments compare cases against absolute standards established as classification thresholds. If a case meets or exceeds standards, it receives one classification. If it falls below standards, it receives a different classification.

Criterion-Referenced Assessment is commonly used in educational testing, where student performances are the cases and grades are classification categories (such as “A”, “B”, *etc.*). Student performance is measured through test questions designed to reflect proficiency levels. The measured performance is compared with standards. Criterion-referenced assessment also is commonly used by government agencies for awarding entitlements and other benefits. Applicants are scored and compared with standards to identify those qualified.

A criterion-referenced assessment is developed through several steps:

- identification of criteria associated with the classification categories;
- development of variables that measure the criteria;
- establishment of threshold standards for variables;
- development of an assessment using the variables;
- development of a scoring system (or procedure) for the assessment;
- measuring cases along the variables;
- scoring cases; and
- classifying cases.

For the classification of “rural residents” and “non-rural residents,” the cases are populations of Alaska residents. In our criterion-referenced assessment, two *primary criteria* and three *ancillary criteria* are identified for distinguishing between rural and non-rural populations. Variables and standards measuring the criteria are defined as the following:

Criterion 1. Country Food Production

Variable: Annual per capita harvest (lbs) of country food.

General Standards:

- “Very High (VH)” >115 lbs (>75% RDA for protein)
- “Moderately High (MH)” 75-114 lbs (50%-74% RDA for protein)
- “Moderately Low (ML)” 40-74 lbs (25%-49% RDA for protein)
- “Very Low (VL)” <39 lbs (<24% RDA for protein)

Criterion 2. Sparsely-Populated, Open Country

Variable: Weighted population in a 30-mile standard area.

General Standards:

- “Yes (Y)” <100 people/100 sq mi
- “No (N)” >100 people/100 sq mi

Criterion 3. Other Extensive Land Uses

Variable: Regular employment in commercial fisheries, forestry, etc.

General Standards:

“Yes (Y)”

“No (N)”

Criterion 4. Noncommercial Fishery or Hunt Center

Variable: Substantial harvest and distribution of specialty country food products.

General Standards:

“Yes (Y)”

“No (N)”

Criterion 5. Preponderance of Other Rural Features

Variables: (a) Diversity of Resources Used; (b) Diversity of Resources Shared; (c) Country-Oriented Knowledge and Values; (d) Geographic Isolation.

General Standards:

“Yes (Y)”

“No (N)”

Using the above criteria and variables, the rules for categorizing cases as “rural” or “non-rural” with the standards are presented in the following matrix.

**Rules for Classifying "Rural"
and "Non-rural" Populations with Criteria**

		Criterion 2. Sparsely-Populated (Open) Country	
		No	Yes
Criterion 1. Country Food Production	Very Low	"Non-rural"	"Non-rural"
	Moderately Low	"Non-rural"	(a)
	Moderately High	(a)	"Rural"
	Very High	"Rural"	"Rural"

(a) "Non-rural" unless one
other rural feature (Criteria 3, 4, 5)

As shown in the above matrix, the two *primary criteria* are Criterion 1 (Country Food Production) and Criterion 2 (Sparsely-Populated, Open Country). Three additional criteria are used as *ancillary criteria* for categorizing uncertain cases following the application of the primary criteria, if necessary. The threshold standards for categorizing cases are designed to result in the following classification outcomes, based on definitions of rural populations and non-rural populations:

Definitions of "Rural" and "Non-Rural" Populations

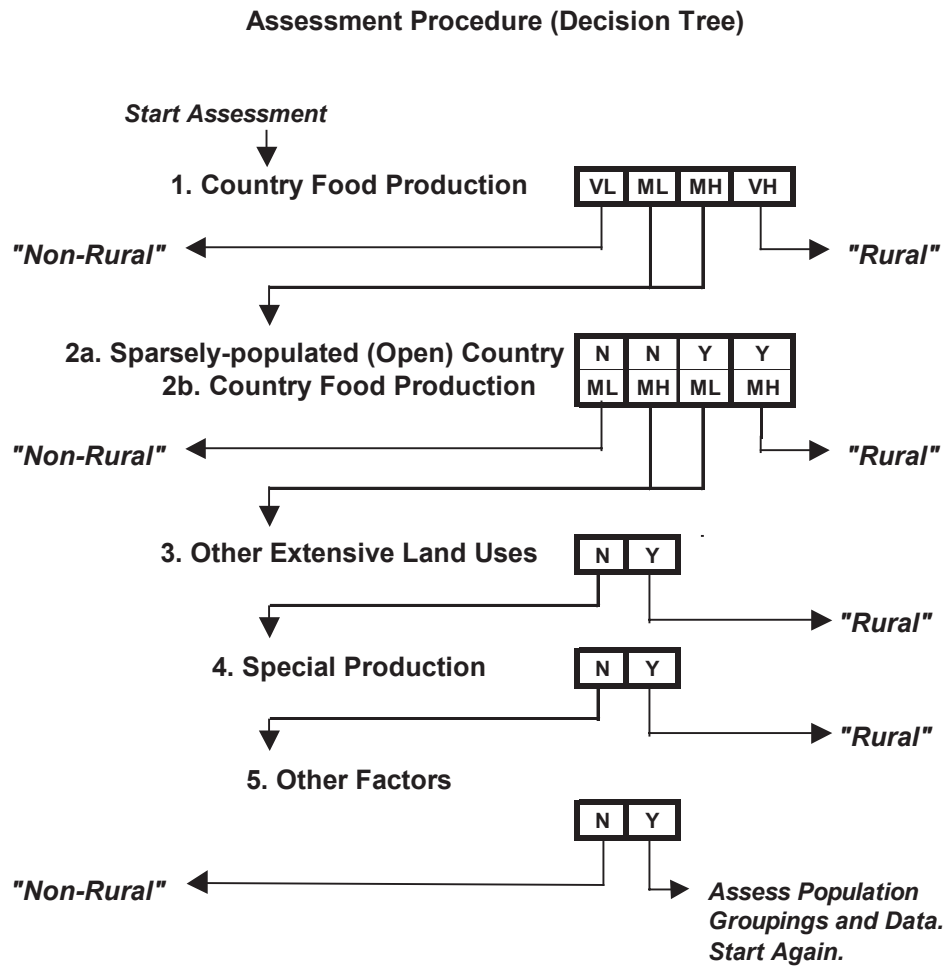
Rural Populations are populations...

1. with very high production of country foods; or
2. with moderately-high production of country foods and in sparsely-populated (open) country; or
3. with moderately-low production of country foods and in sparsely-populated (open) country, or with moderately-high production of country foods and in other than sparsely-populated (open) country, and having one other rural feature:
 - A. regular employment in extensive land uses, or
 - B. a center for a special or distinctive non-commercial fishery or hunt, or
 - C. a preponderance of other factors of a rural nature, such as diversity of resources used, diversity of resources shared, country-oriented knowledge and values, and geographic isolation.

Non-Rural Populations are populations...

1. with very low production of country foods; or
2. with moderately-low production of country foods and in other than sparsely-populated (open) country; or
3. with moderately-low production of country foods and in sparsely-populated (open) country, or with moderately-high production of country foods and in other than sparsely-populated (open) country, and having no other rural feature:
 - A. no regular employment in extensive land uses, and
 - B. no center for a special or distinctive non-commercial fishery or hunt, and
 - C. no preponderance of other factors of a rural nature, such as diversity of resources used, diversity of resources shared, country-oriented knowledge and values, and geographic isolation.

Case populations are assessed with the criteria following a set procedure (decision tree) and series of questions:



Assessment Procedure (Question Series)

1. Is the production of country food "very high," "moderately high," "moderately low," or "very low"?

Measured by the per capita production of country food for local use.

- a. *very high (>115 lbs) = "Rural"*
- b. *moderately high (75-114 lbs) = "Uncertain" (continue)*
- c. *moderately low (40-74 lbs) = "Uncertain" (continue)*
- d. *very low (< 39 lbs) = "Non-rural"*

- 2a. Is the population in "sparsely-populated (open) country"?

Measured by the weighted population in a 30-mile standard area.

AND

- 2b. Is the production of country food "moderately high" or "moderately low"?

Measured by the per capita production of country food for local use.

- a. *yes (< 100 people/100 sq mi) and moderately high (75-114 lbs) = "Rural"*
- b. *yes (< 100 people/100 sq mi) and moderately low (40-75 lbs) = "Uncertain" (continue)*
- c. *no (> 100 people/100 sq mi) and moderately high (75-114 lbs) = "Uncertain" (continue)*
- d. *no (> 100 people/100 sq mi) and moderately low (40-75 lbs) = "Non-rural"*

3. Is the population regularly supported by (employed in) extensive land uses, such as commercial fishing or forestry?

- a. *yes = "Rural"*
- b. *no = "Uncertain" (continue)*

4. Is the population a center for a special or distinctive noncommercial fishery or hunt for non-local distribution and local use?

- a. *yes = "Rural"*
- b. *no = "Uncertain" (continue)*

5. Is there a preponderance of other factors of a rural nature, such as diversity of resources used, diversity of resources shared, country-oriented knowledge and values, and geographic isolation?

- a. *yes = "Rural"*
- b. *no = "Uncertain" (continue)*

6. Are there potential populations or communities whose land use patterns have not been adequately documented?

- a. *yes = Assess population groupings and data. Start again.*
- b. *no = "Non-Rural"*

Application of Criterion-Referenced Assessment to 195 Case Populations (p.1)

No.	Population Name	Characteristics			1. Country Food		2. Open (a) and Country Food (b)		
		People	Density	Harvests	Score	Class	2a Score	2b Score	Class
1	Anchor Point (group)	2,334	74	55	ML	Uncertain	Yes	ML	Uncertain
2	Fritz Creek CDP	1,603	88	72	ML	Uncertain	Yes	ML	Uncertain
3	Hope	137	6	61	ML	Uncertain	Yes	ML	Uncertain
4	Moose Pass (group)	374	16	44	ML	Uncertain	Yes	ML	Uncertain
5	North Fork Road	467	55	71	ML	Uncertain	Yes	ML	Uncertain
6	Trapper Creek	423	19	51	ML	Uncertain	Yes	ML	Uncertain
7	Salcha-Harding	1,128	73	47	ML	Uncertain	Yes	ML	Uncertain
8	Talkeetna (group)	813	19	55	ML	Uncertain	Yes	ML	Uncertain
9	Valdez	4,036	143	103	MH	Uncertain	No	MH	Uncertain
10	Eklutna	4,835	457	42	ML	Uncertain	No	ML	Non-rural
11	Homer	3,946	178	66	ML	Uncertain	No	ML	Non-rural
12	Kasilof (group)	1,639	102	60	ML	Uncertain	No	ML	Non-rural
13	Kenai	6,942	340	60	ML	Uncertain	No	ML	Non-rural
14	Soldotna (group)	14,946	239	42	ML	Uncertain	No	ML	Non-rural
15	Cantwell	222	8	112	MH	Uncertain	Yes	MH	Rural
16	Clam Gulch	173	51	99	MH	Uncertain	Yes	MH	Rural
17	Cooper Landing	369	16	77	MH	Uncertain	Yes	MH	Rural
18	Gakona	215	10	95	MH	Uncertain	Yes	MH	Rural
19	Glennallen	554	23	100	MH	Uncertain	Yes	MH	Rural
20	Nikolaevsk	345	55	89	MH	Uncertain	Yes	MH	Rural
21	Skwentna (group)	148	4	101	MH	Uncertain	Yes	MH	Rural
22	Tazlina	149	16	107	MH	Uncertain	Yes	MH	Rural
23	Voznesenka	327	42	103	MH	Uncertain	Yes	MH	Rural
24	Whittier	182	6	80	MH	Uncertain	Yes	MH	Rural
25	Airport	18,626	2263	18	VL	Non-rural			
26	Avenue Fifteen	12,288	2570	9	VL	Non-rural			
27	Big Lake	2,635	169	20	VL	Non-rural			
28	Campbell Creek	9,245	2805	15	VL	Non-rural			
29	Central Fairbanks	16,788	900	17	VL	Non-rural			
30	Chugiak	4,472	755	37	VL	Non-rural			
31	Coastal Refuge	8,612	1698	21	VL	Non-rural			
32	Delaney Lake	2,917	2873	13	VL	Non-rural			
33	Downtown	1,458	2586	8	VL	Non-rural			
34	Eagle River	20,610	1026	27	VL	Non-rural			
35	Eielson AFB	5,400	341	23	VL	Non-rural			
36	Elmendorf	6,626	1417	18	VL	Non-rural			
37	Fort Richardson	5,470	1281	15	VL	Non-rural			
38	Fort Wainwright	7,381	656	19	VL	Non-rural			
39	Girdwood	2,091	366	18	VL	Non-rural			
40	Glacier View CDP	249	9	36	VL	Non-rural			
41	Houston	1,202	158	12	VL	Non-rural			
42	Juneau City and Boroug	30,711	311	25	VL	Non-rural			
43	Ketchikan	7,922	207	34	VL	Non-rural			
44	Lake Otis	5,275	3071	12	VL	Non-rural			
45	Little Campbell Creek	23,581	2708	15	VL	Non-rural			
46	Lower OMalley-Cambell	12,697	2026	21	VL	Non-rural			
47	Merrill Field	4,128	2914	10	VL	Non-rural			
48	MidFork-RusJack	10,105	2900	12	VL	Non-rural			
49	Midtown	12,687	2982	14	VL	Non-rural			
50	Muldoon	36,961	2573	17	VL	Non-rural			
51	Nikiski	4,327	214	17	VL	Non-rural			
52	North Fairbanks	8,253	292	10	VL	Non-rural			
53	North Pole Area	16,295	739	27	VL	Non-rural			
54	Northeast Fairbanks	4,894	475	33	VL	Non-rural			
55	Northfork	4,324	2913	14	VL	Non-rural			
56	Northwest Fairbanks	5,127	505	16	VL	Non-rural			
57	OMalley	6,000	2118	21	VL	Non-rural			
58	Palmer (group)	15,000	265	27	VL	Non-rural			
59	Rabbit Creek	12,318	1469	23	VL	Non-rural			
60	Russian Jack	4,084	3014	12	VL	Non-rural			
61	Seward (group)	4,670	72	28	VL	Non-rural			
62	Ship Creek	6,727	2556	12	VL	Non-rural			
63	Southwest Fairbanks	17,574	818	19	VL	Non-rural			
64	Spenard	14,939	2589	13	VL	Non-rural			
65	Sutton-Alpine	1,080	121	24	VL	Non-rural			

Application of Criterion-Referenced Assessment Method (p. 2)

No.	Population Name	Characteristics			1. Country Food	
		People	Density	Harvests	Score	Class
66	University	4,633	2964	12	VL	Non-rural
67	Upper OMalley	4,574	1218	22	VL	Non-rural
68	Wasilla (group)	29,618	344	24	VL	Non-rural
69	Willow (group)	2,614	90	23	VL	Non-rural
70	Akhiok	80	3	322	VH	Rural
71	Akutan	713	25	466	VH	Rural
72	Alakanuk	652	23	725	VH	Rural
73	Aleknagik	221	14	379	VH	Rural
74	Allakaket/Alatna	132	5	540	VH	Rural
75	Anderson	367	13	139	VH	Rural
76	Angoon	572	20	224	VH	Rural
77	Anvik	104	4	843	VH	Rural
78	Atka	92	3	439	VH	Rural
79	Barrow	4,581	162	289	VH	Rural
80	Beaver	84	3	457	VH	Rural
81	Bettles-Evansville	71	3	260	VH	Rural
82	Brevig Mission	276	10	579	VH	Rural
83	Chenega Bay	86	3	275	VH	Rural
84	Chickaloon	213	37	224	VH	Rural
85	Chignik Bay	79	3	358	VH	Rural
86	Chignik Lagoon	103	4	211	VH	Rural
87	Chignik Lake	145	5	442	VH	Rural
88	Chistochina	93	4	262	VH	Rural
89	Chitina	123	5	342	VH	Rural
90	Clark's Point	75	9	363	VH	Rural
91	Coffman Cove	199	7	276	VH	Rural
92	Copper Center	362	20	174	VH	Rural
93	Cordova	2,454	87	179	VH	Rural
94	Craig	1,397	49	232	VH	Rural
95	Deering	136	5	672	VH	Rural
96	Dillingham	2,466	89	242	VH	Rural
97	Egegik	116	4	384	VH	Rural
98	Ekwok	130	5	797	VH	Rural
99	Emmonak	767	27	612	VH	Rural
100	False Pass	64	2	413	VH	Rural
101	Fort Yukon	595	21	685	VH	Rural
102	Galena	675	24	368	VH	Rural
103	Golovin	144	5	605	VH	Rural
104	Grayling	194	7	894	VH	Rural
105	Gulkana	88	8	153	VH	Rural
106	Gustavus	429	15	241	VH	Rural
107	Haines	1,811	64	196	VH	Rural
108	Healy	1,000	35	132	VH	Rural
109	Hollis	139	5	169	VH	Rural
110	Holy Cross	227	8	634	VH	Rural
111	Hoonah	860	30	372	VH	Rural
112	Hughes	78	3	567	VH	Rural
113	Huslia	293	10	598	VH	Rural
114	Hydaburg	382	14	384	VH	Rural
115	Hyder	97	3	345	VH	Rural
116	Igiugig	53	2	725	VH	Rural
117	Iliamna	102	4	847	VH	Rural
118	Kake	710	25	179	VH	Rural
119	Kaktovik	293	10	886	VH	Rural
120	Kenny Lake	410	16	136	VH	Rural
121	King Cove	792	28	256	VH	Rural
122	King Salmon	442	16	220	VH	Rural
123	Kivalina	377	13	761	VH	Rural
124	Klawock	854	30	320	VH	Rural
125	Klukwan	139	5	608	VH	Rural
126	Kodiak City	6,334	239	151	VH	Rural
127	Kodiak Road	3,991	109	168	VH	Rural
128	Koliganek	182	6	830	VH	Rural
129	Kotlik	591	21	503	VH	Rural
130	Kotzebue	3,082	109	593	VH	Rural

Application of Criterion-Referenced Assessment Method (p. 3)

No.	Population Name	Characteristics			1. Country Food	
		People	Density	Harvests	Score	Class
131	Kwethluk	713	25	836	VH	Rural
132	Lake Louise	88	3	179	VH	Rural
133	Larsen Bay	115	4	370	VH	Rural
134	Levelock	122	4	884	VH	Rural
135	Manokotak	399	20	384	VH	Rural
136	McGrath	401	14	182	VH	Rural
137	McKinley Park Village	142	5	242	VH	Rural
138	Mentasta Lake	142	7	125	VH	Rural
139	Minto	258	9	585	VH	Rural
140	Mountain Village	755	27	820	VH	Rural
141	Naknek	678	24	188	VH	Rural
142	Nanwalek	177	19	254	VH	Rural
143	Naukati Bay	135	5	242	VH	Rural
144	Nelson Lagoon	83	3	254	VH	Rural
145	New Stuyahok	471	17	700	VH	Rural
146	Newhalen	160	6	747	VH	Rural
147	Nikolai	100	4	450	VH	Rural
148	Ninilchik	772	51	135	VH	Rural
149	Noatak	428	15	461	VH	Rural
150	Northway	95	3	278	VH	Rural
151	Nuiqsut	433	15	742	VH	Rural
152	Nunapitchuk	466	16	802	VH	Rural
153	Old Harbor	237	8	300	VH	Rural
154	Ouzinkie	225	39	264	VH	Rural
155	Pedro Bay	50	2	397	VH	Rural
156	Pelican	163	6	355	VH	Rural
157	Perryville	107	4	394	VH	Rural
158	Petersburg	3,224	114	161	VH	Rural
159	Pilot Point	100	4	384	VH	Rural
160	Point Lay	247	9	890	VH	Rural
161	Port Alexander	81	3	312	VH	Rural
162	Port Alsworth	104	4	361	VH	Rural
163	Port Graham	171	20	253	VH	Rural
164	Port Heiden	119	4	408	VH	Rural
165	Port Lions	256	28	331	VH	Rural
166	Port Protection	63	2	451	VH	Rural
167	Quinhagak	555	20	768	VH	Rural
168	Saint Paul	532	19	267	VH	Rural
169	Sand Point	952	34	256	VH	Rural
170	Saxman	431	162	211	VH	Rural
171	Seldovia	286	31	184	VH	Rural
172	Shageluk	129	5	445	VH	Rural
173	Shishmaref	562	20	794	VH	Rural
174	Sitka	8,835	169	205	VH	Rural
175	Sitka Tribe	2,095	169	350	VH	Rural
176	Slana	124	5	174	VH	Rural
177	South Naknek	137	5	297	VH	Rural
178	Stebbins	547	19	997	VH	Rural
179	Stevens Village	87	3	578	VH	Rural
180	Tanacross	140	10	250	VH	Rural
181	Tanana	308	11	539	VH	Rural
182	Tatitlek	107	10	406	VH	Rural
183	Tenakee Springs	104	4	330	VH	Rural
184	Tetlin	117	6	214	VH	Rural
185	Thorne Bay	557	20	179	VH	Rural
186	Tok	1,393	50	149	VH	Rural
187	Tonsina	92	6	156	VH	Rural
188	Tyonek	193	13	260	VH	Rural
189	Unalaska	4,283	152	195	VH	Rural
190	Wainwright	546	19	751	VH	Rural
191	Wales	152	5	744	VH	Rural
192	Whale Pass	58	2	185	VH	Rural
193	Whitestone Logging Car	116	4	178	VH	Rural
194	Wrangell	2,308	82	167	VH	Rural
195	Yakutat	680	24	386	VH	Rural

Application of Criterion-Referenced Assessment Method (p. 4)

No.	Population Name	3. Extensive Land Use			4. Special Production			5. Other Factors	
		Score	Class		Score	Class		Score	Class
1	Anchor Point (group)	?	Uncertain	→	No	Uncertain	→	?	?
2	Fritz Creek CDP	?	Uncertain		No	Uncertain		?	?
3	Hope	?	Uncertain		No	Uncertain		?	?
4	Moose Pass (group)	?	Uncertain		No	Uncertain		?	?
5	North Fork Road	?	Uncertain		No	Uncertain		?	?
6	Trapper Creek	?	Uncertain		No	Uncertain		?	?
7	Salcha-Harding	?	Uncertain		No	Uncertain		?	?
8	Talkeetna (group)	?	Uncertain		No	Uncertain		?	?
9	Valdez	?	Uncertain		No	Uncertain		?	?

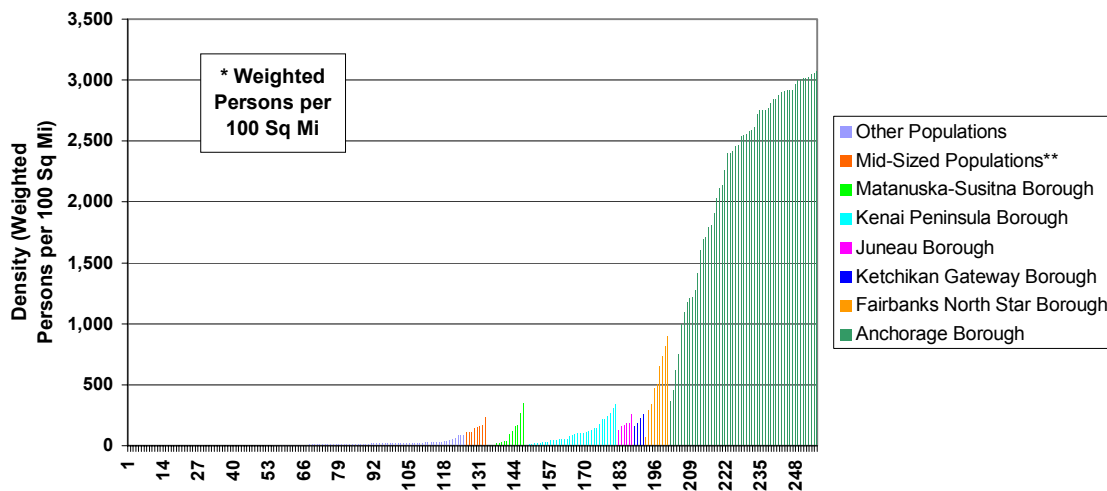
The above criterion-referenced assessment method was applied in the assessment of 195 case populations in Alaska. The case populations comprised the same set as previously described in the *Discriminant Analysis Assessment* section. The outcomes of the criterion-referenced assessment are presented in the table, *Application of Criterion-Referenced Assessment Method*.

As shown in the above table, case populations were initially assessed with the first of the two *primary criteria* – country food production (lbs per person), divided into four threshold levels (“very high,” “moderately high,” “moderately low,” or “very low”). Following classification rules, case populations with “very high” country food production were classified “rural” and cases with “very low” country food production were classified “non-rural.” “Very high” was defined as country food production containing 75% or more of a population’s recommended dietary allowance (RDA) for protein (>115 lbs per person per year), while “very low” was defined as country food production containing less than 25% of a population’s RDA for protein (<39 lbs per person per year). Of 195 case populations, 171 cases (88%) were categorized using the first standard. Of these, 126 cases were classified as “rural” and 45 cases were classified as “non-rural.” Of the 24 uncertain cases, 13 had “moderately low” country food production (25-49% RDA for protein; between 40-74 lbs per person per year) and 11 had “moderately high” country food production (50-74% RDA for protein; between 75-114 lbs per person per year).

In accordance with the established procedure, the 24 unclassified cases were jointly assessed with two primary criteria: *country food production* and *sparsely-populated, open country*. “Sparsely-populated, open country” was defined as < 100 persons (weighted) per 100 sq miles in a local commons, as measured by the variable DENS30 (this variable was described in the previous *Discriminant Analysis* section). Conceptually, the standard for sparsely-populated, open country is equivalent to a community of about 2,500 people whose neighboring 30-mile standard area contains a small number of other people. The standard 30-mile area (a circle with a radius of 30 miles) comprises 2,826 sq miles. Therefore, 2,500 people / 2,826 sq mi = 88 people per 100 sq mi. An additional small number of people (326 weighted people more) within the community's 30-mile standard area raises the total to 100 people per 100 sq mi. This represents a fairly strict, though reasonable, standard for “sparsely-populated, open country.” Such a cutting point can be directly related to the *rural presumption standard* in current federal subsistence regulations (populations no greater than 2,500 people). The measure represents a

“presumed” rural community (<2,500 people) surrounded by a relatively unpopulated open area. The location of this cut-point can be seen in Fig. 18, which shows density values (DENS30) of 255 case populations in Alaska. The same information set is shown in Fig. 10, but as log transformed data.

Fig. 18. Densities* of Standard Use Areas of 255 Populations Grouped by Borough



Following classification rules, case populations with “sparsely-populated, open country” and “moderately high” country food production were categorized as “rural.” Cases without “sparsely-populated, open country” and “moderately low” country food production were categorized as “non-rural.” As shown in the assessment table above, 10 of the 24 unclassified cases were categorized as “rural” and five were classified as “non-rural” by this procedure. There remained only nine cases unclassified – eight cases with moderately low country food production in sparsely-populated, open country, and one case (Valdez) with moderately high country food production and without sparsely-populated, open country. The unclassified cases included six populations along roads near a rural/non-rural fringe (Anchor Point, Trapper Creek, Talkeetna, North Fork Road, Fritz Creek, and Salcha-Harding), one mid-sized, geographically-separate population (Valdez), and two small communities (Hope and Moose Pass).

Following the established procedure, the nine unclassified cases would be assessed using three *ancillary criteria* (*other extensive land uses, noncommercial fishery or hunt center, and a preponderance of other rural features*). A case meeting one or more of the ancillary criteria would be classified as “rural.” A case having no other rural feature would be classified as “non-rural.” The ancillary criteria likely would be examined by the Federal Subsistence Board using existing data sets, case method materials, and public comment.

In the evaluation of *Other Extensive Land Uses* (Criterion 3), COMMFISH (a variable in the PACK database) measures the percentage of households with members involved in commercial fishing, derived from the CPDB. Statistics on commercial fishing permits (number and percentage fished) and annual commercial fish harvests (numbers and lbs of fish) by Alaska communities are available through the Alaska Commercial Fisheries Entry Commission or the Alaska Department of Community and Economic Development. These measures can be used to assess the relative involvement of community residents in commercial fisheries. An illustration of this information is presented the following table, *Example of Information on Commercial Fisheries by Case Population*. In addition, economic summaries of case communities are generally available from the Alaska Department of Community and Economic Development's website (www.dced.state.ak.us/cbd/commdb/CF_COMDB.htm). For instance, for Valdez, the DCED summary states, "In 2000, gross fishing revenues of residents exceeded \$1.6 million... three fish processing plants operate in Valdez, including Peter Pan and Seahawk Seafoods." A recent history of commercial fishing in the Pacific Gulf of Alaska, states "Valdez became a participant in [Pacific Gulf] commercial fisheries during this period [1975-1995], but oil transport was and remains the primary economic engine of this community" (Fall et al 2001:52ff). Annual fish sales ranged between \$1.0-2.5 million during the 1990s (Fall et al 2001:56). The per capita income from commercial fishing at Valdez was the lowest among surveyed Pacific Gulf communities (Fall et al 2001:131). Such information can be used by the federal board to assess if a community is regularly supported by (employed in) extensive land uses.

Example of Information on Commercial Fisheries by Case Population				
Case Population	Residents (2000)	Residents with Commercial Fishing Permits (DCED)	Percentage of Residents with Commercial Fishing Permits (DCED)	Percentage of Households with Members Employed in Commercial Fisheries (CPDB)
Anchor Point	1,845	80	4%	--
Fritz Creek CDP	1,603	13	1%	15.0% (1999)
Hope	137	2	1%	2.9% (1990)
Moose Pass	206	2	1%	--
North Fork Road	467	--	--	10.2% (1998)
Salcha-Harding	854	4	0.5%	--
Talkeetna	772	10	1%	0% (1985)
Trapper Creek	423	7	2%	0% (1985)
Valdez	4,336	42	1%	9.6% (1992)

Noncommercial Fishery or Hunt Center (Criterion 4) refers to communities with special noncommercial fisheries or hunts that exist in a few locations in Alaska. Examples of noncommercial fishery centers include the herring roe-on-hemlock fishery in Sitka Sound (Robert F. Schroeder and Matt Kookesh, *Subsistence Harvest of Herring Eggs in Sitka Sound*, Technical Paper No. 173, Division of Subsistence, ADF&G, 1990) and the hooligan fishery for oil production on the Chilkat and Chilkoot rivers in southeast Alaska (Martha F. Betts, *The Subsistence Eulachon Fishery of the Chilkat and Chilkoot Rivers*,

Southeast Alaska, Technical Paper No. 213, Division of Subsistence, ADF&G, 1994). During the 1990s, production of herring roe-on-hemlock and hooligan oil for regional distribution occurred in a few special locations like these in southeast Alaska. Barrow provides an example of a major hunt center for bowhead whale (Rosita Worl and Charles W. Smythe, *Barrow: A Decade of Modernization*, Technical Report No. 125, Minerals Management Service, Alaska OCS Socioeconomic Studies Program, 1986). The rural or non-rural designation of a fishery center or hunt center potentially impacts the production and distribution of specialty country food products over a wider area. This potential impact warrants the consideration of this feature as a rural criterion.

A Preponderance of Other Rural Features (Criterion 5) is a general criterion that allows for a wide number of factors to be considered in assessing a borderline case, such as *Diversity of Resources Used*, *Diversity of Resources Shared*, *Country-Orientated Knowledge and Values*, and *Geographic Isolation*. The *Diversity of Resources Used* in a community is indicated by variables such as SPECOUNT, USECOUNT, NUSED50, HRVCOUNT, and NHARV50 in the PACK Database (see Appendix A). The *Diversity of Resources Shared* in a community is indicated by variables such as PCTGVALL, PCTRCALL, PCTGVSLM, PCTRCSLM, PCTGVMLM, PCTRCLML, GIVECOUNT, RECCOIUNT, PCTGIVEN, PCTRECVD, NGIV25, NREC25, PCTGIV1, PCTREC1, PCTGIV2, and PCTREC2. *Country-Oriented Knowledge and Values* are qualitative variables. They may be indicated by measures such as INDIGNDX (a measure of the extent that traditional food items are used in a population) and RRFISH1 and RRFISH2 (the percentage of a community's fish harvested with rod and reel gear) in the PACK Database. For particular case populations, predominant value orientations may be pertinent to its rural or non-rural classification. Areas where country foods are primarily derived by sport fishing and sport hunting may be more similar to non-rural Alaska areas than rural Alaska areas. *Geographic Isolation* refers to the place of a community in relation to other populations. COLAVG, the cost of imported food in an area, indicates whether a community is toward the periphery of commercial food trade networks. Lower food costs are usually found in population centers, while higher food costs are typical of peripheral populations.

Because measures for ancillary variables may be unavailable for many Alaska populations, rigorous statistical comparisons of cases may not be possible for certain ancillary measures. Instead, ancillary variables may be assessed through a procedure called *Case Method Assessment*. In Case Method Assessment, detailed information is gathered on a set of case populations, commonly across a range of key variables. The purpose is to describe a few cases in substantial detail so as to understand relationships among variables specific to that set of cases. The case histories provide background to contemporary information. Both qualitative and quantitative information is usually analyzed together. The information on cases is presented as a narrative accompanied with charts, graphs, maps, and other exhibits.

Case Method Assessment allows for a fuller understanding of land and resource use patterns in Alaska than can be portrayed in statistical approaches focused on a few variables. The depth of information clarifies relationships among a range of factors.

The case analysis helps to explain why relationships are or are not seen. Case Method is particularly useful for understanding populations that may be relatively unique. Populations that deviate from normative patterns may not be adequately accounted for in general statistical models. The Case Method approach allows for information to be collected and analyzed on such unique populations, so that “exceptions to the rule” are understood and reasonably assessed.

Case examples of resource use systems are commonly presented in the form of regional profiles, community profiles, or household cases. The technical report series from the *Division of Subsistence, Alaska Department of Fish and Game* is a major source of case materials

(www.state.ak.us/local/akpages/FISH.GAME/subsist/geninfo/publctns/subabs.htm). The community profile website of the *Alaska Department of Community and Economic Development* is another source of information

(www.dced.state.ak.us/cbd/commdb/CF_COMDB.htm). Information provided by members of regional advisory groups and expert stakeholders is a third potential source. Collectively, these materials offer important information for making decisions on subsistence uses in Alaska, whatever rural assessment approach is utilized.

Comparison of Outcomes

The two assessment methods (*Discriminant Analysis Assessment* and *Criterion-Referenced Assessment*) applied similar criteria in substantially different fashions. The one approach involved multivariate modeling with interval scaled data, where the contribution of variables to defining groups emerged through relatively complex inductive statistics. The other approach was a “top-down” deductive method, applying relatively simple standards (“very high,” “moderately high,” “moderately low,” “very low”; “yes,” “no”) defined through a mix of logic, reason, and empirical evidence.

While differing in approach, each method produced similar classification outcomes. Out of 195 case populations, there were no cases with divergent classifications. That is, no case was classified “rural” by one method and “non-rural” by the other method. For 182 of 195 cases (93%), the two methods provided identical classifications. For the remaining 13 cases, the methods presented some differences in the degree of certainty of classifications, as shown in the following summary, *Tentative or Uncertain Classifications*. The discriminant analysis gave tentative classifications to four rural cases and six non-rural cases; of these ten cases, the criterion-referenced assessment gave uncertain classifications to one of the rural cases and five of the non-rural cases. The criterion-referenced assessment gave uncertain classifications to three cases that received certain classifications by the discriminant analysis assessment (two non-rural cases and one rural case), as shown in the summary.

Tentative or Uncertain Case Classifications		
Population	Discriminant Analysis	Criterion-Referenced Assessment
Anchor Point (group)	Non-rural	Uncertain
Salcha-Harding	Non-rural	Uncertain
Fritz Creek CDP	Tentative Non-rural	Uncertain
Glacier View CDP	Tentative Non-rural	Non-rural
Moose Pass (group)	Tentative Non-rural	Uncertain
North Fork Road	Tentative Non-rural	Uncertain
Talkeetna (group)	Tentative Non-rural	Uncertain
Trapper Creek	Tentative Non-rural	Uncertain
Clam Gulch	Tentative Rural	Rural
Kodiak City	Tentative Rural	Rural
Nikolaevsk	Tentative Rural	Rural
Valdez	Tentative Rural	Uncertain
Hope	Rural	Uncertain

The substantial similarity in classification outcomes suggests the two methodologies are making similar differentiations between rural populations and non-rural populations in Alaska, using the rural/non-rural criteria. Using two very different approaches (inductive statistics and deductive reasoning), rural and non-rural groups are distinguishable by country food production levels (a type of extensive land use) and sparsely-populated, open country (measured by weighted population within standard areas). The consistency in the groupings of case populations between the two approaches provides cross validation of the methods and factors. Because of the consistency of outcomes, one may feel more secure in the choice of one or the other methods in classifying case populations.

The classifications of the two methodologies can be compared with the current rural and non-rural classifications in federal regulations. As described above, the federal findings were made using a third, substantially-different methodology. The federal findings were made by the Federal Subsistence Board applying information pertaining to a set of factors, including but not limited to *use of fish and wildlife, development and diversity of the economy, community infrastructure, transportation, and educational institutions*. There are initial presumption levels based on population size, with communities less than 2,500 people presumed “rural” and greater than 7,000 presumed “non-rural.” The findings are made in a public process, commonly including substantial testimony of regional advisory councils and other stakeholders. A qualitative assessment is made considering the weight of information, rather than a quantitative assessment. The federal approach resembles approaches used by the Alaska State Joint Board of Fisheries and Game in making rural and nonsubsistence area determinations, as described above.

The outcomes from the two tested methodologies (*Discriminant Analysis Assessment* and *Criterion-Referenced Assessment*) are substantially similar to the current rural and non-rural classifications in federal regulations. With a few exceptions, the communities and areas designated as “rural” and “non-rural” in federal regulations are similarly designated under the two quantitative approaches. The outcomes of the *Discriminant Analysis Assessment* are consistent with current federal findings except for Valdez (“tentative

rural”), Clam Gulch (“tentative rural”), and three populations on the fringe of the Wasilla area -- Glacier View, Talkeetna, and Trapper Creek (“tentative non-rural”). The classifications of these places by the *Discriminant Analysis Assessment* were tentative, indicating that the populations’ scores were greater than one standard deviation from their closest group. Such tentative classifications might be reasonably changed in light of information from additional *ancillary factors*. The outcomes of the *Criterion-Referenced Assessment* applying the two primary criteria were consistent with current federal findings, except for Glacier View (“non-rural”). However, nine places were left unclassified in the test of the *Criterion-Referenced Assessment* – Anchor Point, Salcha-Harding, Fritz Creek, Moose Pass, North Fork Road, Talkeetna, Trapper Creek, Valdez, and Hope – pending the application of *ancillary factors* through a case method approach. The final classifications using the ancillary criteria may or may not be entirely consistent with current federal findings.

The comparison of outcomes can be viewed as a test of validity. If one assumes that the current federal classifications are substantially correct for subsistence management purposes, their consistency with the outcomes of the two tested methodologies can be interpreted as a validation of the new, quantitative approaches. That is, the two new methodologies appear to be making distinctions among Alaska populations that are similar to those made by the current Federal Subsistence Board procedure. The three methodologies appear to be finding consistent contrasts between rural and non-rural groups in Alaska.

Consistency in outcomes is not too surprising. The two new methodologies are applying information on country food harvests and demography similar to information used by the Federal Subsistence Program in making the current rural classifications. So one might anticipate some similarity in outcomes through the use of similar assessment factors. Further, the *Discriminant Analysis Assessment* employs federal and state classifications as initial guides to rural and non-rural groups in Alaska, which assists locating statistical breaking points between groups. Similarly, the density standard used in the *Criterion-Referenced Assessment* is linked to presumption levels found in federal regulation. These similar features would lead to some convergence of outcomes.

However, the two new methodologies also apply new information and substantially different approaches in reaching its outcomes. The country food harvest information for Alaska’s large population centers is essentially new. Demographic information is applied in an essentially new fashion through the density criterion. The information on populations and geographic areas, derived from the 2000 federal census, is more recent than information used in past rural/non-rural findings. A new approach for aggregating and disaggregating populations in certain road-connected areas is used. Further, the application of the information in the two new quantitative assessments differs substantially from the more qualitative approaches used by the federal and state programs. Considering these kinds of differences in information and assessment approaches, it also would not have been surprising if the two new methodologies had produced substantially different outcomes compared with past assessments.